

1           **SIGNAL SEPARATION METHOD AND APPARATUS FOR RESTORING**  
2           **ORIGINAL SIGNAL FROM OBSERVED DATA**

3    **FIELD OF THE INVENTION**

4    The present invention relates to signal separation and  
5    processing. It more particularly relates to signal  
6    separation of an original signal when multiple mixed signals  
7    are observed.

8    **BACKGROUND OF THE INVENTION**

9    In some cases, a technique whereby, when multiple mixed  
10   signals are observed by multiple sensors, original signals  
11   are decided only by using observed signals are more useful  
12   than conventional noise reduction methods, and some  
13   extensions of the application field can be expected. These  
14   fields include, for example, speech enhancement for the  
15   reduction of unwanted acoustics during speech recognition,  
16   digital communication demodulation in a complex signal  
17   environment such as QAM (Quadrature Amplitude Modulation), a  
18   medical signal restoration for the extraction of necessary  
19   organ information, and a data analysis method whereby an  
20   independent component (factor) hidden in statistical data  
21   can be extracted.

1 Figure 1 depicts a conceptual diagram showing signal  
2 separation problems for separating original signals,  
3 assuming only the mutually statistical independence of  
4 signals when multiple signals are observed in a mixed state.  
5 Figure 1 is formulated as follows. First, assume that there  
6 are m signals of scalar values  $s_1(t)$ , ...,  $s_m(t)$  for each  
7 index t, which are mutually statistically independent and  
8 have zero mean. Figure 1 shows two signal sources  $s_1$  and  $s_2$   
9 by way of example. For this, n linear weighted sums  $x_1(t)$ ,  
10 ...,  $x_n(t)$  are to be observed, which is expressed as follows  
11 and observed by an observation apparatus.

12 [Equation 1]

$$\underline{x}(t) = \underline{A} \underline{s}(t)$$

14 where each element is represented as follows.

15 [Equation 2]

$$\underline{x}(t) = [x_1(t) x_2(t) \cdots x_n(t)]^T$$

$$\underline{s}(t) = [s_1(t) s_2(t) \cdots s_m(t)]^T$$

18 where it is assumed to be  $n \geq m$ . Furthermore, assuming that n  
19  $m \times n$  mixing matrix is  $\underline{A}$ , which is to be a full rank matrix,  
20 i.e., a matrix where an inverse matrix of  $m \times m$  matrix  $\underline{A}^H \underline{A}$   
21 exists. Hereinafter, a lowercase letter with an underline  
22 represents a vector, an uppercase letter with an underline  
23 represents a matrix, a subscript T represents transposition,  
24 and a subscript H represents Hermitian conjugate (i.e.,  
25 conjugate transposition).

1 A problem of estimating a separation matrix  $\underline{W}$  for obtaining  
2 a separation signal  $\underline{y}(t)$  from this observed signal  $\underline{x}(t)$  is a  
3 so-called signal separation problem. That is, when  
4 obtaining a separation signal  $\underline{y}(t) = \underline{W}^H \times \underline{x}(t)$  using a signal  
5 separation apparatus from an observed signal  $\underline{x}(t)$ , which was  
6 observed by an observation apparatus shown in Figure 1,  
7 estimation of a separation matrix  $\underline{W}$  becomes a problem.

8 Next, a summary of the concept for estimating a separation  
9 matrix  $\underline{W}$  will be described. Assuming that a multivariable  
10 probability density function of a signal vector serving as  
11 observed signal  $\underline{x}(t)$  is  $p_u(\underline{u})$  and a probability density  
12 function for each element of the vector is  $p_i(u_i)$ , a mutual  
13 information of an observed vector is represented by the  
14 following Kullback-Leibler divergence.

15 [Equation 3]

$$I(\underline{u}) = \int p_u(\underline{u}) \log \left( \frac{p_u(\underline{u})}{\prod_{i=1}^n p_i(u_i)} \right) d\underline{u}$$

17 where the mutual information is always positive and when it  
18 is zero shows that the elements of each signal vector are  
19 independent. In fact, if the signal vector elements are  
20 independent each other, the density function of the signal  
21 vector is represented by the following equation, so that the  
22 above equation becomes zero.

1 [Equation 4]

2 
$$p_u(\underline{u}) = \prod_{i=1}^n p_i(u_i)$$

3 Therefore, one of the rationales of signal separation  
4 technique is that the original signals are able to be  
5 restored from the mixed observed signals by finding a  
6 transformation matrix that minimizes the mutual information  
7 of signal vectors for observed signal vectors.

8 However, as the probability distribution of original signals  
9 is practically unknown, the mutual information can not be  
10 made directly to be an object of minimization operation.  
11 Therefore, the signal separation is often performed by  
12 optimizing a valuation amount that is equal or approximately  
13 equal to the mutual information. For example, Reference 1  
14 "International Journal of Neural Systems", vol. 8, Nos. 5 &  
15 6, pp. 661-678, October/December 1997, describes that a  
16 mutual information is able to be minimized if finding a  
17 transformation matrix  $\underline{W}$  that optimizes the sum of the  
18 fourth-order cumulants with a zero time delay for each  
19 original signal (i.e., maximizing if the kurtosis is  
20 positive or minimizing if the kurtosis is negative), on the  
21 condition that the observed signals have a kurtosis with the  
22 same sign, a covariance matrix is bounded, whitening has  
23 been performed, and a separation matrix  $\underline{W}$  is a unitary  
24 matrix (i.e.,  $\underline{W}^H \underline{W} = \underline{I}$  (unit matrix)). Note that the kurtosis  
25 refers to a numeric obtained by the following calculation

1 for an observed signal  $u_i$ .

2 [Equation 5]

$$3 \quad E\{u_i^4\} - 3[E\{u_i^2\}]^2$$

4 where  $E[\cdot]$  represents an expectation operation. The  
5 whitening means making signal vectors uncorrelated each  
6 other to make the variance 1, the fourth-order cumulant is a  
7 statistic represented by the following equation.

8 [Equation 6]

$$\begin{aligned} 9 \quad c_4(k_1, k_2, k_3) = & E\{u_i(t)u_i(t+k_1)u_i(t+k_2)u_i(t+k_3)\} \\ 10 & - E\{u_i(t)u_i(t+k_1)\}E\{u_i(t+k_2)u_i(t+k_3)\} \\ 11 & - E\{u_i(t)u_i(t+k_2)\}E\{u_i(t+k_1)u_i(t+k_3)\} \\ 12 & - E\{u_i(t)u_i(t+k_3)\}E\{u_i(t+k_1)u_i(t+k_2)\} \end{aligned}$$

13 The zero time delay means that  $k_1$ ,  $k_2$  and  $k_3$  are zero in the  
14 above equation.

15 However, generally, as a load of calculation is heavy when  
16 calculating high order statistics such as cumulants, a  
17 technique is employed such as calculating and approximating  
18 another information amount equivalent to a mutual  
19 information or minimizing a cost function equivalent to what  
20 optimizes the sum of cumulants by introducing a nonlinear  
21 function that can approximate the fourth-order cumulants.  
22 U.S. patent No. 5706402 discloses a method for finding a  
23 separation matrix by the gradient method using an  
24 unsupervised learning algorithm that optimizes output  
25 entropy instead of minimization of mutual information.

1    Though Reference 2 (Signal Processing, vol. 24, No. 1, pp.  
2    1-10, July 1991) does not manifest mutual information and  
3    cumulants, it discloses a method for using an approach  
4    similar to it, wherein a square of the residual that results  
5    from subtracting a linear sum of estimated signals from the  
6    observed signal is made to be a cost function and finding a  
7    separation filter that minimizes the cost function by the  
8    gradient method. Moreover, Japanese Unexamined Patent  
9    Publication No. 2000-97758 discloses a method for improving  
10   the convergence by normalizing updated amounts of the above  
11   method.

12   Reference 3 (IEEE Transactions on Signal Processing, vol.  
13   44, No. 12, pp. 3017-3030, December 1996) proposes an  
14   estimation method, wherein a nonlinear function that  
15   approximately finds the fourth-order cumulants is  
16   introduced, then updated amounts to optimize the cost  
17   function in an adaptive algorithm based on that nonlinear  
18   function are determined based on the relative gradient.  
19   This technique improves the convergence speed of the  
20   conventional adaptive algorithm, which uses a gradient of  
21   the cost function as the updated amount, and which is  
22   equivalent to the natural gradient that may be introduced  
23   from information geometric considerations.

24   Stability in the convergence process of the separation  
25   matrix is important when restoring signals not in a steady  
26   state. In fact, in a series of gradient methods described  
27   above, it is often the case that the relation between the

1 convergence speed and the stability is an inverse  
2 proportion. Thus, U.S. Patent No. 5999956 uses a method  
3 that adds a module for reducing the effect on the estimation  
4 process even when there is a big change of power between  
5 estimated signals, and outputting stable results, in  
6 addition to a signal estimation module and a separation  
7 coefficient estimation module in order to achieve a stable  
8 convergence.

9 Furthermore, Reference 4 (International Journal of Neural  
10 Systems, vol. 8, No. 5 & 6, pp. 601-612, October/December  
11 1997) derives an adaptive algorithm based on the least  
12 squares method instead of the gradient method, when  
13 optimizing a cost function that introduced nonlinear  
14 function. Using this approach, as a step-size is not  
15 determined by a user like in the gradient method and what is  
16 optimal is determined automatically, the convergence speed  
17 is enhanced and the stability is achieved under a given  
18 condition.

19 Like the technique of Reference 4 above, within the  
20 framework of the least squares method, it has been  
21 considered that a fast and appropriate convergence is often  
22 achieved, since a step-size is calculated to be optimal  
23 under the cost function. However, there is not necessarily  
24 the conformance between the situation where the signal  
25 separation is required and the format of the cost function  
26 which the above prior art techniques including the gradient  
27 method have been employed, so that there is a case where it  
28 seems not to be best even when using the framework of the

1 least squares method.

2 For example, for a portable information device, it is  
3 assumed that signal observation apparatuses are close to  
4 each other because a large area can not be obtained for the  
5 installation of apparatuses. At this time, it is easily  
6 assumed that the original signals can be mixed at a similar  
7 ratio by the observation apparatuses. When this mixing  
8 ratio is represented as a matrix element, the elements in  
9 each column (or each row) have substantially the same value.

10 In such a case, as the condition number of the mixing  
11 matrix becomes large, the perturbation in the estimation  
12 process of the separation matrix would have large effects in  
13 estimates. Note that the condition number refers to an  
14 amount defined by  $\|\underline{Z}\| \cdot \|\underline{Z}^{-1}\|$  using some norm  $\|\cdot\|$  for a  
15 matrix  $\underline{Z}$ , where  $\underline{Z}^{-1}$  represents an inverse matrix of a  
16 matrix  $\underline{Z}$ .

17 Therefore, in the conventional format of the cost function,  
18 much time is spent for obtaining normal estimates when the  
19 perturbation is large, which is likely to be a problem.  
20 Further, it is another problem that when the condition  
21 number is not large, the convergence speed becomes slower  
22 than the conventional cost function in the stage where  
23 errors still remain in the estimation process.

#### 24 SUMMARY OF THE INVENTION

25 To resolve the above technical problems, it is one aspect of



1 the present invention to stably extract, using only a small  
2 number of calculation steps, the original signals from  
3 multiple mixed signals that are observed.

4 It is another aspect of the present invention to provide a  
5 fast convergent, least squares type calculation method,  
6 based on the cost function, that when the perturbation  
7 affecting an estimated value is large, this change is  
8 reflected by proper weighting for signal separation.

#### 9 BRIEF DESCRIPTION OF THE DRAWINGS

10 These and other aspects, features, and advantages of the  
11 present invention will become apparent upon further  
12 consideration of the following detailed description of the  
13 invention when read in conjunction with the drawing figures,  
14 in which:

15 Fig. 1 depicts a conceptual diagram showing signal  
16 separation problems.

17 Fig. 2 shows an example of a block diagram showing a first  
18 algorithm for estimating a separation matrix that minimizes  
19 a cost function.

20 Fig. 3 shows an example of an overall configuration of a  
21 second algorithm.

22 Fig. 4 shows an example of an estimation filtering of  $\underline{w}_i(t)$   
23 in a second algorithm;

1 Fig. 5 shows an example of a flow of processing from data  
2 reading to data outputting according to the present  
3 invention.  
4 Fig. 6 shows an example of a conceptual diagram showing a  
5 degree of consideration of an error in the cost function.  
6 Fig. 7 depicts the results of convergence when conducting  
7 the independent trial ten times and taking an average of  
8 them in the experiment of separation.  
9 Figs. 8A to 8D are diagrams showing the original signals of  
10 real speech.  
11 Figs. 9A to 9D are diagrams showing mixed speech signals.  
12 Figs. 10A to 10D are diagrams showing the separation results  
13 obtained by this embodiment.

#### 14 DESCRIPTION OF THE SYMBOLS

15 21: Nonlinear function  
16 22: Calculation of error signal  $\underline{e}(t)$   
17 23: Update of  $\underline{W}(t)$   
18 24: Unitarization operation  
19 25: Calculation of  $\underline{h}(t)$   
20 26: Calculation of  $\underline{g}(t)$  and  $\xi$   
21 27: Calculation of  $\underline{P}(t)$   
22 31: Estimation filtering of  $\underline{w}_1(t)$   
23 32: Estimation filtering of  $\underline{w}_2(t)$   
24 33: Estimation filtering of  $\underline{w}_m(t)$   
25 41: Nonlinear function

- 1           42: Calculation of error signal  $\underline{e}_i(t)$   
2           43: Update of  $\underline{w}_i(t)$   
3           45: Update of  $\underline{x}_{i+1}(t)$   
4           46: Calculation of  $\xi$   
5           47: Calculation of  $\underline{d}_i(t)$

6    DESCRIPTION OF THE INVENTION

7    To achieve these aspects, according to the present  
8    invention, an adaptive algorithm is employed for introducing  
9    a function, such as an exponential type function having a  
10   monotonously increasing characteristic, as a cost function  
11   that provides effects equivalent to the minimization of the  
12   mutual information for observed signals, and for minimizing  
13   (optimizing) the cost function relative to a signal  
14   separation matrix. The results acquired from the  
15   optimization of the exponential type function are equivalent  
16   to those acquired through the optimization of the H-infinity  
17   norm and the solution of the two-person zero-sum game in the  
18   game theory, that is, the solution of so-called MinMax  
19   strategy. Thus, the signal separation matrix can also be  
20   estimated using the adaptive algorithm that employs the  
21   H-infinity norm and MinMax strategy. That is, according to  
22   the present invention, a signal separation method for  
23   restoring an original signal from observed data (observed  
24   signals), obtained by observing multiple mixed signals,  
25   comprises the steps of: estimating, from the observed data,

1 a separation matrix using an adaptive filter that suppresses  
2 the H-infinity norm concerning the separation matrix until  
3 the H-infinity norm is to equal to or smaller than a  
4 provided scalar value; and restoring the original signal by  
5 multiplying the separation matrix by the observed data.

6 The signal separation method of this invention further  
7 comprises the steps of: selecting, for the observed data, a  
8 specific separation matrix from among multiple separation  
9 matrixes based on MinMax strategy in game theory; and  
10 restoring an original signal by multiplying the selected  
11 separation matrix by the observed data.

12 That is, the signal separation method of this invention can  
13 be an optimization method for, based on the MinMax strategy  
14 of game theory, selecting as a solution a separation matrix  
15 to which is output a minimum error value selected from among  
16 maximum error values that are output to various separation  
17 matrixes.

18 Further, a signal separation method for estimating and  
19 restoring an original signal from observed data obtained by  
20 observing multiple mixed signals, which include the original  
21 signal, comprises the steps of: introducing, for the  
22 observed data, a cost function based on a function having a  
23 monotonously increasing characteristic; estimating a  
24 separation matrix using an adaptive filter that optimizes  
25 the cost function; and estimating and restoring the original  
26 signal by multiplying the separation matrix by the observed  
27 data.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

For the estimate of the separation matrix, an adaptive filter is employed for minimizing the cost function for the separation matrix, and the cost function to be introduced is an exponential type function. With these functions, even when large perturbation is applied to the estimated value during the process for estimating the separation matrix, the time used to obtain a correct estimated value can be reduced.

A signal separation method for this invention comprises the steps of: reading observed signals; subtracting the average of the observed signals and performing zero averaging for the observed signals; whitening the observed signals obtained by zero averaging; separating the whitened observed signals based on a cost function that has a monotonously increasing characteristic; and performing, as a post processing, inverse whitening for the obtained observed signals. A non-linear function to be used in the cost function employed for the separation is changed, based on the kurtosis of the observed signal.

A signal processing apparatus according to the invention comprises: input means, for receiving observed data obtained by observing multiple mixed signals, which include an original signal; separation matrix estimation means, for estimating, for the observed data, a separation matrix using adaptive filtering for suppressing the H- infinity norm concerning the separation matrix until the H-infinity norm is to equal to or smaller than a provided scalar value; and estimation/restoration means, for estimating and restoring

1 the original signal by multiplying the separation matrix by  
2 the observed data.

3 A signal processing apparatus for this invention comprises:  
4 input means, for receiving observed data obtained by  
5 observing multiple mixed signals, which include an original  
6 signal; selection means, for employing, for the observed  
7 data, the MinMax strategy in game theory to select, for  
8 example, a separation matrix that outputs a minimum error  
9 selected from among maximum errors output from separation  
10 matrixes; and estimation/restoration means, for estimating  
11 and restoring an original signal by multiplying the  
12 separation matrix by the observed data.

13 A signal processing apparatus for this invention comprises:  
14 separation matrix estimation means, for estimating, for  
15 input observed data, a separation matrix by using an  
16 adaptive filter with optimizing a cost function that is  
17 based on a function, such as an exponential type function,  
18 having a monotonously increasing characteristic, e.g., for  
19 minimizing the cost function for the separation matrix; and  
20 estimation/restoration means, for estimating and restoring  
21 an original signal by multiplying the separation matrix by  
22 the observed data.

23 From another viewpoint, according to the invention, a signal  
24 processing apparatus, for separating an original signal from  
25 an input observed signal and outputting the original signal,  
26 comprises: a non-linear function unit, for performing a  
27 non-linear function for an input observed signal and a  
28 separation matrix estimated during the previous cycle; an

1 error signal calculator, for calculating an error signal  
2 based on the value obtained by the non-linear function unit,  
3 the separation matrix estimated during the previous cycle,  
4 and an observed signal at the present time; and a separation  
5 matrix update unit, for updating the separation matrix  
6 estimated at that time based on the error signal, so that  
7 error evaluation is weighted by the cost function having the  
8 monotonously increasing characteristic.

9 It is advantageous, because accuracy is improved, that the  
10 signal processing apparatus further comprise a unitarization  
11 operator (a quadrature operator for a real number signal)  
12 for ensuring the transform to a unitary matrix (a quadrature  
13 matrix for a real number signal) of the estimated separation  
14 matrix that has been updated by the separation matrix update  
15 unit at that time.

16 Furthermore, according to the invention, a signal processing  
17 apparatus, for separating multiple mixed speech signals that  
18 are observed when, for example, multiple speakers are  
19 speaking simultaneously, comprises: input means, for  
20 receiving mixed speech data; separation matrix estimation  
21 means, for estimating a separation matrix, for the mixed  
22 speech data, using an adaptive filter with optimizing a cost  
23 function that is based on a function having a monotonously  
24 increasing characteristic; and separation/extraction means,  
25 for separating and extracting the speech signals from the  
26 mixed speech data by multiplying the separation matrix by  
27 the mixed speech data. According to this aspect, the speech  
28 of a single, targeted speaker can be accurately extracted,

1 regardless of whether other, included signals are those  
2 produced by the speech of other speakers or are those  
3 produced by environmental noise. This provides an effective  
4 pre-processing, such as is required for speech recognition.

5 According to the invention, a signal processing apparatus  
6 for separating an artifact, such as the mixing of potential  
7 changes, from an observed bio-signal that, at the least, is  
8 either a signal observed using magnetoencephalography (MEG)  
9 or electroencephalography (EEG), comprises: input means, for  
10 receiving observed data containing the artifact in the  
11 observed bio-signal; separation matrix estimation means, for  
12 estimating a separation matrix for the observed data, using  
13 an adaptive filter with optimizing a cost function that is  
14 based on a function having a monotonously increasing  
15 characteristic; and separation/extraction means for  
16 separating and extracting the observed bio-signal from the  
17 observed data by multiplying the separation matrix by the  
18 observed data. According to these inventions, even when the  
19 mixing process is unknown, it can be expected that a brain  
20 active potential signal will be accurately extracted,  
21 without removing the original brain waves.

22 According to another aspect of the invention, a signal  
23 processing apparatus for extracting, from economic  
24 statistical data, a fluctuation element that is hidden  
25 during the observation process, comprises: input means, for  
26 receiving economic statistical data; separation matrix  
27 estimation means, for estimating a separation matrix for the  
28 economic statistical data using an adaptive filter with



1 optimizing a cost function that is based on a function  
2 having a monotonously increasing characteristic; and  
3 separation/extraction means, for separating and extracting  
4 the fluctuation element from the economic statistical data  
5 by multiplying the separation matrix by the economic  
6 statistical data.

7 The economic statistical data, received by the input means,  
8 is management data that can be considered as the data that  
9 consists of the overall trend and individual factors  
10 synthesized by an unknown mixing matrix. Thus, this  
11 invention can be used for the extraction of the main factor  
12 that affects cash flow, for example.

13 The economic statistical data are stock price fluctuation  
14 data that are observed as a set, and the fluctuation element  
15 that is separated and extracted by the separation/extraction  
16 means is a stock price trend for an independent component  
17 that can be applied for portfolio return prediction. With  
18 this arrangement, for the investment division determinations  
19 performed in the financial field, the main factor affecting  
20 a fluctuation in stock prices can be analyzed more  
21 accurately, and the time used to estimate the main factor  
22 can be reduced considerably.

23 According to the invention, a mobile terminal device, for  
24 receiving, from a base station for code division multiple  
25 access, observed data that include the spread information to  
26 other users, and for extracting a local user signal from the  
27 observed data, comprises: input means, for receiving  
28 observed data from the base station; separation matrix

estimation means, for estimating a separation matrix for the observed data using an adaptive filter with optimizing a cost function that is further based on a function having a monotonously increasing characteristic; and separation/extraction means, for separating and extracting a user signal from the observed data by multiplying the separation matrix by the observed data. According to this aspect, an affect due to fading, or due to a signal from another user whose spread code is unknown, can be modeled as the coefficient of a mixing matrix, and the target user signal can be separated only from the received observed data.

A storage medium is provided on which a computer stores a computer-readable program that permits the computer to perform: a process for introducing, for observed data obtained by observing multiple mixed signals, including the original signal, a cost function that is based on a function having a monotonously increasing characteristic; a process for estimating a separation matrix using an adaptive filter that optimizes the cost function; and a process for estimating and restoring the original signal by multiplying the separation matrix by the observed data.

This program also permits a computer to perform: a process for estimating, for observed data obtained by observing multiple mixed signals, which include an original signal, a separation matrix through adaptive filtering for suppressing the H-infinity norm concerning the separation matrix until the H-infinity norm is to equal to or smaller than a provided scalar value; and a process for restoring the

1 original signal by multiplying the separation matrix by the  
2 observed data.

3 The program also permits a computer to perform: a process  
4 for selecting from observed data obtained by observing  
5 multiple mixed signals, which include an original signal, a  
6 specific separation matrix from among separation matrix  
7 candidates based on the MinMax strategy in game theory; and  
8 a process for restoring an original signal by multiplying  
9 the separation matrix by the observed data.

10 The storage medium can be, for example, a CD-ROM, and the  
11 program can be read by the CD-ROM reader of a computer and  
12 stored on the hard disk of the computer, for example, and  
13 executed.

#### 14 Advantageous Embodiment

15 An advantageous embodiment will now be described in detail  
16 while referring to the accompanying drawings. A signal  
17 separation method according to this embodiment can be  
18 applied for a field for reducing an artifact from an  
19 observed bio-signal produced by magnetoencephalography (MEG)  
20 or electroencephalography (EEG). This method can also be  
21 applied for speech enhancement for the reduction of unwanted  
22 acoustics during a speech recognition process, or for signal  
23 separation or interference reduction during digital  
24 communication, such as QAM (Quadrature Amplitude  
25 Modulation). Further, the method can be applied for the  
26 reduction of an interference set up by down-link signals  
27 that serve as links from a base station to individual mobile  
28 terminals for CDMA (Code Division Multiple Access), or for

1 the extraction of a mobile terminal user signal. Or the  
2 method can be used as a data analysis method for the  
3 extraction of an important fluctuation element (factor),  
4 such as a stock price fluctuation, that is hidden from  
5 economic statistical data during the observation process, or  
6 for portfolio management in the financial field. That is,  
7 the signal separation method of this embodiment can be  
8 widely applied for problems that require the separation of  
9 an original signal from general, one-dimensional mixed  
10 signals that are observed, regardless of whether these  
11 signals are digital observed signals or analog observed  
12 signals, or are complex signals or real signals. The method  
13 of the invention can be applied not only for the above  
14 examples, but also for derivative forms of these examples.  
15 Furthermore, in the above fields, the signal separation  
16 technique of the invention can be supplied by a computer  
17 that includes various memories and controllers and a display  
18 device, and also by a special apparatus or terminal.

19 Next, a signal separation method according to the present  
20 invention will be described in detail.

21 First, by introducing a nonlinear function that can  
22 approximate a fourth-order cumulant into a cost function and  
23 making its minimization have a meaning equivalent to  
24 minimization of mutual information, a scheme to estimate a  
25 separation matrix represented as follows is considered.

26 [Equation 7]

27 Separation Matrix  $\widehat{W}$

1 Conventionally, a cost function used in such a scheme is the  
2 following format.

3 [Equation 8]

$$4 \quad E\{\|\underline{x} - \hat{\underline{W}}\underline{f}(\hat{\underline{W}}^H \underline{x})\|^2\}$$

5 where  $\underline{f}(\cdot)$  is a vector that has a nonlinear function such as  
6  $\tanh(\cdot)$  as an element.

7 According to the present invention, a signal separation  
8 method is proposed that minimizes the following type of cost  
9 function in terms of the above separation matrix. That is,

10 [Equation 9]

$$11 \quad J_1(\hat{\underline{W}}) = \gamma^2 \ln E\{\exp(\gamma^{-2} \|\underline{x} - \hat{\underline{W}}\underline{f}(\hat{\underline{W}}^H \underline{x})\|^2)\}$$

12 Then, assuming that a difference between an initial true  
13 value of the separation matrix and an estimated initial  
14 value, and a noise generated in the estimation process are  
15 zero mean and decorrelated, and an estimated noise is a  
16 white random variable having unit variance. A white random  
17 variable means that  $v_i$  and  $v_j$  for any  $i$  and  $j$  ( $i \neq j$ ) are  
18 an independent stochastic variable in a variable series  $[v_k]$ .

19 In the description below, an index ( $t$ ) may be added instead  
20 of a hat symbol that indicates an estimate of the separation  
21 matrix, as long as there is no need to distinguish between a  
22 true value and an estimated value.

1 Minimization of the above cost function means minimization  
2 of the following equation in view of that a logarithmic  
3 function  $\ln(\cdot)$  is a monotonously increasing function.

4 [Equation 10]

$$5 \quad E\{\exp(\gamma^{-2} \|\underline{x} - \hat{\underline{W}}\underline{f}(\hat{\underline{W}}^H \underline{x})\|^2)\}$$

6 In addition, though various proposals have been provided as  
7 to what kind of nonlinear function should be used, it is  
8 common to use threshold processing for changing a function  
9 to be used depending on whether the kurtosis of the observed  
10 signals is positive or negative, thereby selecting a  
11 function that enables an appropriate approximation of high  
12 order cumulants. Other various types of forms are also  
13 conceivable. Functions to be used include  $\tanh(u)$ ,  $u^3$ , or  
14  $u - \tanh(u)$ , etc. For example,  $u - \tanh(u)$  is used when the  
15 kurtosis is positive, while  $\tanh(u)$  is used when the kurtosis  
16 is negative.

17 Figure 2 depicts a block diagram showing a first algorithm  
18 for estimating a separation matrix that minimizes a cost  
19 function. The algorithm shown in Figure 2 is represented as  
20 follows.

21 [Equation 11]

$$22 \quad \underline{y}(t) = \underline{f}(\underline{W}^H(t-1)\underline{x}(t))$$

$$23 \quad \underline{h}(t) = \underline{P}(t-1)\underline{y}(t)$$

$$24 \quad \underline{g}(t) = \underline{h}(t) / [\beta + \underline{y}^H(t)\underline{h}(t)]$$

$$\begin{aligned}
1 \quad & a = 1 - \gamma^{-2} \\
2 \quad & \xi = \{\beta + \underline{y}^H(t)\underline{h}(t)\} / \{\beta + a\underline{y}^H(t)\underline{h}(t)\} \\
3 \quad & \underline{P}(t) = \frac{1}{\beta} \{\underline{P}(t-1) - a\xi \underline{g}(t)\underline{h}^H(t)\} \\
4 \quad & \underline{e}(t) = \underline{x}(t) - \underline{W}(t-1)\underline{y}(t) \\
5 \quad & \underline{W}(t) = \underline{W}(t-1) + \underline{e}(t)\underline{g}^H(t)
\end{aligned}$$

6 Figure 2 depicts a flow of main algorithm by a nonlinear  
7 function 21, calculation of an error signal  $\underline{e}(t)$  22, update  
8 of  $\underline{W}(t)$  23, and a unitarization operation 24, wherein  
9 various amounts necessary for this main algorithm are  
10 calculated by each unit including calculation of  $\underline{h}(t)$  25,  
11 calculation of  $\underline{g}(t)$  and  $\xi$  26, and calculation of  $\underline{P}(t)$  27.  
12  $z^{-1}$  is a delay for timing adjustment. As shown in Figure 2  
13 and the above algorithm, first an observed signal  $\underline{x}(t)$  is  
14 input, and then  $\underline{y}(t)$  is calculated by the nonlinear function  
15 21 that is determined depending on whether the kurtosis is  
16 positive or negative. This nonlinear function 21 is  
17 constituted such that the nonlinear function 21 is operated  
18 on a product of an estimated separation matrix  $\underline{W}(t-1)$  that  
19 was estimated at a previous time (i.e., previous cycle) and  
20 an observed signal  $\underline{x}(t)$  at a present time. Thereafter,  $\underline{y}(t)$   
21 is input to the calculation of an error signal  $\underline{e}(t)$  22,  
22 where a difference is calculated between the observed signal  
23  $\underline{x}(t)$  at that time and a product of the estimated separation  
24 matrix  $\underline{W}(t-1)$  at a previous time (i.e., previous cycle) and

1 the calculated  $\underline{y}(t)$ . Next, the update of the separation  
 2 matrix  $\underline{W}(t)$  at that time is performed in a block 23. In  
 3 this update of  $\underline{W}(t)$  23,  $\underline{W}(t)$  is updated by multiplying  
 4  $\underline{W}(t-1)$  estimated at the previous time by the error signal  
 5  $\underline{e}(t)$  and  $\underline{g}^H(t)$ . The unitarization operation 24 is a step for  
 6 assuring a unitary matrix after the update of  $\underline{W}(t)$ , where a  
 7 given operation is advantageously performed for improving  
 8 the accuracy. This is expressed as follows.

9 [Equation 12]

$$\underline{W}(t) = \underline{W}(t) \{ (\underline{W}(t)^H \underline{W}(t))^{-1} \}^{1/2}$$

11 In the calculation of  $\underline{h}(t)$  25,  $\underline{h}(t)$  is calculated from  $\underline{P}(t-1)$   
 12 at the previous time and  $\underline{y}(t)$ . In the calculation of  $\underline{g}(t)$   
 13 and  $\xi$  26, each value is calculated using a forgetting factor  
 14  $\beta$ . This forgetting factor  $\beta$  is a constant of  $0 < \beta \leq 1$ , which  
 15 may be usually 1, however, when a matrix  $\underline{W}$  to be estimated  
 16 varies with time, this forgetting factor  $\beta$  is set to smaller  
 17 than 1 such as 0.99 or 0.98 to enable tracking for this  
 18 varying. In the calculation of  $\underline{P}(t)$  27,  $\underline{P}(t)$  is calculated,  
 19 which is an inverse matrix of a covariance matrix of  $\underline{y}(t)$ ,  
 20 from  $\underline{h}(t)$  calculated in the calculation of  $\underline{h}(t)$  25, input  $a$ ,  
 21 the forgetting factor  $\beta$ , and  $\underline{g}(t)$  and  $\xi$  calculated in the  
 22 calculation of  $\underline{g}(t)$  and  $\xi$  26. Note that initial values  $\underline{P}(0)$   
 23 and  $\underline{W}(0)$  are arbitrarily selected for these algorithms.



1 Now assuming that an observed signal  $\underline{x}(t)$  undergoes  
2 preprocessing for transforming into a signal having a zero  
3 mean and performing whitening, before the above estimation  
4 processing of a separation matrix. This preprocessing is  
5 performed on the observed signal  $\underline{x}(t)$  and needs no  
6 information of the original signal. Whitening means that  
7 each element of the observed signal vector is made  
8 uncorrelated to make the variance 1, a technique of which  
9 may not be specified and possible using eigenvalue  
10 decomposition or principal component analysis, for example.

11 In this way, the signal separation method according to the  
12 present invention is based on optimizing a cost function  $J_1$   
13 based on an exponential function in terms of a separation  
14 matrix. Therefore, from this point of view, a similar  
15 result would be obtained by using other algorithms.  
16 Figure 3 and Figure 4 depict a configuration of the second  
17 algorithm. Figure 3 depicts an overall configuration of a  
18 second algorithm, while Figure 4 depicts an estimation  
19 filtering of  $\underline{w}_i(t)$ . This is what the first algorithm is  
20 transformed by deflation. In this second algorithm,  
21 optimization is performed by estimating a column vector one  
22 by one to reduce the order in turn, as shown in Figure 3.  
23 For example, estimation filtering of  $\underline{w}_1(t)$  is performed on  
24 the observed signal  $\underline{x}_1(t)$  in block 31, then estimation  
25 filtering of  $\underline{w}_2(t)$  is performed on the observed signal  $\underline{x}_2(t)$   
26 in block 32, and in the same manner, estimation filtering of

1  $\underline{w}_m(t)$  is performed on the observed signal  $\underline{x}_m(t)$  in block 33,  
 2 thereby finally deriving the estimated separation matrix  
 3  $\underline{W}(t)$ .

4 That is, in this second algorithm, an operation shown in  
 5 Figure 4 is repeated for each  $\underline{x}_i(t) = \underline{x}(t)$ ,  $i=1, \dots, m$ .

6 The algorithm shown in Figure 4 is represented as follows.

7 [Equation 13]

$$y_i(t) = f(\underline{w}_i^H(t-1)\underline{x}_i(t))$$

$$a = 1 - \gamma^{-2}$$

$$\underline{e}_i(t) = \underline{x}_i(t) - \underline{w}_i(t-1)y_i(t)$$

$$\xi = \frac{\beta d_i(t-1) + y_i(t)^* y_i(t)}{\beta d_i(t-1) + a y_i(t)^* y_i(t)}$$

$$\frac{1}{d_i(t)} = \frac{1}{\beta} \left\{ \frac{1}{d_i(t-1)} - a \xi \frac{y_i(t)^* y_i(t)}{d_i(t-1) \{\beta d_i(t-1) + y_i(t)^* y_i(t)\}} \right\}$$

$$\underline{w}_i(t) = \underline{w}_i(t-1) + \underline{e}_i(t) [y_i(t)^* / \{\beta d_i(t-1) + y_i(t)^* y_i(t)\}]$$

$$\underline{x}_{i+1}(t) = \underline{x}_i(t) - \underline{w}_i(t) y_i(t)$$

15 where  $d_i(0)$  is any initial value, a subscript \* represents  
 16 conjugate.  $f(\cdot)$  represents a nonlinear function such as  $\tanh(\cdot)$   
 17 .

18 In Figure 4, a flow of main algorithm is represented by a  
 19 nonlinear function 41, calculation of an error signal  $\underline{e}_i(t)$   
 20 42, update of  $\underline{w}_i(t)$  43, and update of  $\underline{x}_{i+1}(t)$  45, wherein  
 21 various amounts necessary for this main algorithm are  
 22 calculated by each unit including calculation of  $\xi$  46 and

1 calculation of  $\underline{d}_i(t)$  47.  $z^{-1}$  is a delay for timing  
2 adjustment. The accuracy is enhanced by performing a  
3 unitarization operation when each  $\underline{w}_i(t)$  is found or when all  
4  $\underline{w}_i(t)$  are found. For example, in the case of a real vector,  
5 Gram-Schmidt orthogonalization process may be applied,  
6 however, there is no need to restrict a method.

7 Figure 5 is a flowchart showing the processing from the  
8 reading of data to the output of data in this embodiment.  
9 First, data for an observed signal  $\underline{x}(t)$  is read (step 101),  
10 and selection is performed, dependent on whether an  
11 established non-linear function should be employed for an  
12 algorithm or a function should be changed in accordance with  
13 the kurtosis (steps 102 to 104). That is, it is determined  
14 whether a predetermined function is used or not (step 102),  
15 if so, the predetermined function is set (step 103),  
16 otherwise the function is set depending on the determination  
17 of kurtosis (step 104). Next, zero averaging of the  
18 observed signal  $\underline{x}(t)$  is performed, that is, an average of  
19 the observed signal  $\underline{x}(t)$  is subtracted to make the average  
20 be zero (step 105), then whitening of the observed data is  
21 performed (step 106). Next, from the beginning of data to  
22 the end, or from the beginning to a predetermined data  
23 point, separation processing is recursively performed, which  
24 is based on an exponential function type of cost function  
25 according to the present invention (step 107), then the  
26 inverse whitening is performed as a post-processing (step  
27 108), and finally the output result is obtained (step 109).

1 As the separation processing, the configuration based on the  
2 first algorithm shown in Figure 2 or the configuration based  
3 on the second algorithm shown in Figure 3 and Figure 4 may  
4 be used.

5 Next, characteristic of a signal separation method according  
6 to the present invention will be described. As previously  
7 stated, a cost function that has been used before is the one  
8 that has the following form based on  $H^2$  norm.

9 [Equation 14]

$$J_2(\hat{W}) = E\{\|\underline{x} - \hat{W}f(\hat{W}^H \underline{x})\|^2\}$$

11 Such type of cost function estimates a separation matrix  
12 with equal consideration weight even when some kind of  
13 perturbation such as an estimation error is added to the  
14 estimated vector (represented by the following equation) or  
15 when the convergence is obtained.

16 [Equation 15]

$$\hat{\underline{x}} = \hat{W}f(\hat{W}^H \underline{x})$$

18 On the other hand, according to an exponential function type  
19 of cost function of the present invention, as the error  
20 increases, the consideration weight involved in cost  
21 function increases, so that the convergence is accelerated,  
22 while when close to convergence, a stable estimation is  
23 expected so as not to accelerate the unnecessary update.

1 Figure 6 depicts a conceptual diagram showing a degree of  
2 consideration of an error in the cost function. The  
3 horizontal axis shows the value of the error, while the  
4 vertical axis shows a degree of consideration. An  
5 approximate straight line in the drawing indicates the prior  
6 scheme, while a curved line with a downward convex shape  
7 indicates the scheme of the present invention. According to  
8 the present invention, in the initial stage of estimation  
9 where an estimation error is large, a large update is  
10 performed with an optimal update amount, while in the stage  
11 where the estimation proceeds, an appropriate update for  
12 fine tuning is performed steadily, so that consequently the  
13 estimated result with high-precision is obtained by a small  
14 number of steps.

15 An example of the extension of a difference from the  
16 conventional case is an example wherein a large area can not  
17 be obtained to install a signal observation apparatus, such  
18 as a portable information device. At this time, components  
19 of each row or column of a mixing matrix  $\underline{A}$  have the same  
20 value, wherein the condition number of the mixing matrix  $\underline{A}$   
21 becomes large. As the condition number becomes large, the  
22 perturbation of the estimation error or the like has a great  
23 influence, thus the prior scheme is likely to need extra  
24 steps for convergence even when a small perturbation occurs.

25 On the other hand, according to the present scheme, as the  
26 consideration weight in the cost function becomes large when  
27 the large perturbation occurs, an effect equivalent to  
28 making a correction large is possible in the update equation

1 in the algorithm. As a result, according to the present  
2 scheme, it is expected to accelerate the convergence. Of  
3 course, according to the present invention, as is evident  
4 from the above consideration that the convergence is  
5 accelerated compared with the prior scheme when the  
6 condition number is not large, so that the separation matrix  
7 with high-precision is expected to be estimated more  
8 quickly. That is, according to the present scheme, an  
9 update amount that tries to suppress an estimation error as  
10 much as possible is improved under a more appropriate cost  
11 function than the prior one.  
12 Further, as is apparent from the described algorithm, the  
13 present scheme is considered to have a structure of adaptive  
14 filtering, so that even when a mixing matrix A is  
15 time-varying, e.g., the mixing state is changed during the  
16 observation, signal separation with tracking the  
17 time-variation can be performed.

18 According to the above explanation, provided that an  
19 exponential function type of cost function is used and its  
20 calculation method is to optimize the cost function  $J_1$ , the  
21 present invention is not limited to the above-mentioned  
22 form. A similar algorithm may be derived from a cost  
23 function that is based on an approximate expression where an  
24 exponential function is expressed in a polynomial expansion,  
25 or a cost function that is based on a function having a  
26 similar monotonously increasing characteristic. Now, a cost  
27 function will be described, which has the same meaning as an  
28 exponential function type of cost function. That is, by

deriving an estimation algorithm of a separation matrix based on two cost functions described below, an algorithm similar to the ones shown in Figure 2, Figure 3 and Figure 4 can be derived.

[Equation 16]

$$J_{11}(\hat{W}) = \sup_{\underline{W}_0} \frac{E\{\|\underline{W}\underline{f}(\underline{W}^H \underline{x}) - \hat{\underline{W}}\underline{f}(\hat{\underline{W}}^H \underline{x})\|^2\}}{(\underline{W}_0 - \hat{\underline{W}}_0)^H \underline{\Pi}_0^{-1} (\underline{W}_0 - \hat{\underline{W}}_0) + E\{\|\underline{x} - \underline{W}\underline{f}(\underline{W}^H \underline{x})\|^2\}} < \gamma^2$$

$$J_{12}(\hat{W}) = \min_{\hat{\underline{W}}} \max_{\underline{W}_0} [E\{\|\underline{W}\underline{f}(\underline{W}^H \underline{x}) - \hat{\underline{W}}\underline{f}(\hat{\underline{W}}^H \underline{x})\|^2\} - \gamma^2 \{(\underline{W}_0 - \hat{\underline{W}}_0)^H \underline{\Pi}_0^{-1} (\underline{W}_0 - \hat{\underline{W}}_0) + E\{\|\underline{x} - \underline{W}\underline{f}(\underline{W}^H \underline{x})\|^2\}\}]$$

where  $\underline{\Pi}_0$  is a positive-definite matrix representing a degree of uncertainty for an initial value  $\underline{W}_0$ ,  $\underline{W}$  is a true value of a separation matrix. Note that a positive-definite matrix means a matrix  $\underline{M}$  where  $\underline{v}^H \underline{M} \underline{v}$  becomes positive for all non-zero vectors  $\underline{v}$ .

$J_{11}$  is an H-infinity norm and the format of this cost function means that no matter how large an estimation error exists, a method can be derived which estimates a separation matrix, wherein the upper bound of  $H^2$  norm of estimation errors is less than or equal to a predetermined scalar quantity. A cost function  $J_{12}$  means that a method for optimization can be derived in a strategy using the MinMax

1 theorem of the game theory, wherein a separation matrix that  
2 generates a minimum error is selected as a solution, from  
3 among the candidates of a maximal value of errors that arise  
4 for various separation matrices. What these cost functions  
5 mean is consistent with the features of the signal  
6 separation method derived from the exponential function type  
7 of cost functions described above.

8 Next, the results of the experiment will be described when  
9 applying a signal separation method according to the present  
10 invention, wherein three kinds of synthetic signals were  
11 prepared to conduct an experiment of their separation.  
12 These synthetic signals are as follows.

13 [Equation 17]

14  $r_1 = \sin(2\pi 60t)$

15  $r_2 = \text{sign}(\sin 2\pi 500t)$

16  $r_3$ : random variables uniformly distributed between  $[0, 1]$

17 where the sampling frequency was 10 kHz,  $\text{sign}(\cdot)$  represents a  
18 function that outputs a sign of  $(\cdot)$ .

19 The methods of the aforementioned Reference 3 and Reference  
20 4 were used as comparative objects. The method of Reference  
21 3 is the one that is based on a relative gradient of which  
22 convergence is said to be fast among gradient methods, while  
23 the method of Reference 4 is a conventional least squares  
24 type of algorithm. As a comparative object, the algorithm  
25 of the present invention shown in Figure 2 is employed. It



1 was assumed that a mixing matrix is generated from random  
2 variables that are uniformly distributed between [0, 1] and  
3 that the condition numbers are over 2000.

4 Figure 7 depicts the results of convergence when conducting  
5 the independent trial ten times and taking an average of  
6 them in the experiment of separation. The horizontal axis  
7 represents the number of iterations, while the vertical axis  
8 represents an index value of convergence, wherein the  
9 characteristics of convergence by using synthetic signals  
10 are shown corresponding to when using a technique of  
11 Reference 3, a technique of Reference 4, and a technique of  
12 the present invention. As the characteristic index of  
13 convergence at this time, the following index used in the  
14 Reference 4 is employed, which indicates that the precision  
15 of estimation becomes higher when approaching zero.

16 [Equation 18]

$$C = \sqrt{\sum_i \left( \sum_j \frac{|p_{ij}|^2}{\max_k |p_{ik}|^2} - 1 \right) + \sum_j \left( \sum_i \frac{|p_{ij}|^2}{\max_k |p_{kj}|^2} - 1 \right)}$$

18 where  $p_{ij}$  represents (i, j) element of a matrix that is  
19 obtained as a product of an estimated separation matrix and  
20 a mixing matrix  $\underline{A}$ . As the technique of Reference 3  
21 requires that a user determines the update step-size, here a  
22 step-size that achieves the fastest convergence is used  
23 among the ones that achieve the convergence at the same  
24 level of stability as the present invention. However, as is  
25 evident from Figure 7, the convergence speed of the

1 technique of Reference 3 is slower than the technique of the  
2 present invention. On the other hand, although a step-size  
3 can be automatically set to an optimal one according to the  
4 technique of Reference 4 and the technique of the present  
5 invention, it is evident that the technique of the present  
6 invention achieves a faster convergence than the technique  
7 of Reference 4.

8 As the result of examination of the number of FLOPS  
9 (Floating Operations Per Second), i.e., the number of  
10 floating operations that can be processed per second for  
11 respective techniques in the above experiment, it proved  
12 that the number of FLOPS for one step is nearly the same for  
13 the present invention and the Reference 4. Therefore, it is  
14 understood that the total number of FLOPS to convergence is  
15 smaller for the technique of the present invention. On the  
16 other hand, comparing the Reference 3 and the present  
17 invention, the number of FLOPS required for one step is  
18 smaller for the technique of Reference 3, however, the  
19 simulation proved that the total number of FLOPS used for  
20 satisfying convergence is reduced to less than 2/3 by the  
21 present invention. When using the algorithms shown in  
22 Figure 3 and Figure 4, nearly equal results can be obtained.  
23 Moreover, when not making the condition number of mixing  
24 matrix A an unfavorable condition like this simulation, a  
25 similar result can be obtained.

26 Figs. 8 to 10 are diagrams for explaining speech signal  
27 separation results as an example for implementing real

1 signal separation. Figs. 8A to 8D are diagrams showing the  
2 original signals of real speech. Figs. 9A to 9D are  
3 diagrams showing mixed speech signals, which are obtained by  
4 mixing the signals in Figs. 8A to 8D using a mixing matrix  
5 generated by employing uniformly distributed random  
6 variables. Figs. 10A to 10D are diagrams showing the  
7 separation results obtained in this embodiment. The  
8 horizontal axis in each graph represents a sample number,  
9 and the vertical axis represents an amplitude. In Figure  
10 8A, the real speech for /n/, /i/, /N,g/, /e/, /N/, /t/, /o/,  
11 /h/, /a/ (ningen-towa; "a human being is" in Japanese) is  
12 shown. In Figure 10A, the separation results of mixed  
13 speech signals are shown, and it is apparent that the same  
14 signals as in Figure 8A are obtained and that the original  
15 signal was stably separated from multiple signals by the  
16 method of the invention.

17 In Figure 8B, the real speech of /b/, /i/, /my/, /o,u/, /d/,  
18 /e/, /f,u/, /k/, /u/, /z/, /a/, /ts/ (bimyo-de fukuzatsu;  
19 "delicate and complicated" in Japanese) is shown. In Figure  
20 10B, the separation results provided by the mixed speech  
21 signals, and the real speech in Figure 8B are separated.  
22 Furthermore, in Figure 8C, the real speech for /i/, /k/,  
23 /i/, /m/, /o/, /n/, /o/, /d/, /e/, /a/, /r/, /u/ (ikimono-de  
24 aru; "a living thing" in Japanese) is shown. In Figure 8D,  
25 the real speech for /f,u/, /k/, /u/, /z/, /a/, /ts/, /u/,  
26 /n/, /a/, /i/, /k/, /i/, /m/, /o/, /n/, /o/ (fukuzatsu-na  
27 ikimono; "a complicated living thing" in Japanese) is shown.

28 In Figs. 10C and 10D, the speech signals separated for this  
29 embodiment are shown, and it is apparent that the same

1 amplitudes shown in Figs. 8C and 8D are obtained. As is  
2 described above, when multiple speech signals (other speech  
3 signals) mix with a target real speech signal, the method of  
4 the invention can be used to stably separate the target real  
5 speech signal.

6 An explanation will now be given for examples (1) to (3)  
7 illustrating the application of the signal separation method  
8 of this invention to other fields.

9 **(1) Reduction of the artifact from an observed bio-signal**  
10 **provided by magnetoencephalography (MEG) or**  
11 **electroencephalography (EEG)**

12 The signal separation method of the invention is one type of  
13 independent component analysis, and can be expected to be  
14 used as a method for separating a brain active potential  
15 signal from artifacts in a MEG or EEG record. The artifacts  
16 consist of active potential of heart, blinking, the  
17 movements of eyeballs or changes in the myoelectric  
18 potential, an electrical/magnetic disturbance due to  
19 environment, and the malfunction of a sensor. These should  
20 be separated from a signal waveform that represents the  
21 brain activity. However, when multi-channel electrodes are  
22 attached to the surface of a head for the observation of  
23 bio-signals, signals representing brain activity and the  
24 above artifacts are observed at the same time.

25 Conventionally, artifacts are reduced to a degree as the  
26 result of a time and frequency assumption; however, brain

1 wave signals also tend to be removed. In this embodiment,  
2 even when the mixing process is unknown, only a  
3 comparatively short observation time is used to separate and  
4 extract an original signal, and a more accurate extraction  
5 of brain active potential signals can be expected.

6 **(2) Reduction of the interference produced by a down-link**  
7 **signal, which serves as a link from a base station to**  
8 **an individual mobile terminal, at the time of a code**  
9 **division multiple access (CDMA)**

10 The CDMA is a multiple access technique based on spread  
11 spectrum. In this embodiment, a user spreads an information  
12 spectrum by using spread codes, and multiple users  
13 communicate with each other over the same frequency band.  
14 Thus, the spread information for one different user is  
15 interference noise for other users. In addition,  
16 communication is also affected by fading due to multi-paths.

17 The detection and the estimation of delays in transmission  
18 code are important in order to reduce the influence of  
19 deteriorating factors, and the conventional reception  
20 technique includes the detection of time delays using a  
21 matched filter or a maximum likelihood method. The  
22 conventional method for providing an efficient operation can  
23 satisfactorily cope with a location, such as a base station,  
24 that is equipped with a satisfactory signal processing  
25 facility. However, it is anticipated that it will be more  
26 difficult when a complicated signal processing apparatus,  
27 which increases the accuracy of the conventional system, is

1 attached to a mobile terminal for which a reduction in size  
2 and in the power consumed is desired. Thus, when the method  
3 of this embodiment is employed, whereby an affect produced  
4 by fading or a signal from another user having an unknown  
5 spread code is modeled as the coefficient for a mixing  
6 matrix, and whereby only a signal from a local user is  
7 separated from the received signal, the signal can be  
8 received more accurately than by the conventional method.

9 The mobile terminal device to which the embodiment is  
10 applied includes various functions, such as communication  
11 means, control means and display means, that are employed  
12 for a common portable telephone or PDA (Personal Digital  
13 Assistants) known to those skilled in the art. Thus, a  
14 detailed explanation for this need not be given.

15 **(3) Usage of a data analysis method for extracting an**  
16 **important fluctuation element (factor), such as a stock**  
17 **price fluctuation, that can not be detected from**  
18 **economic statistical data, and for portfolio management**  
19 **in the financial field.**

20 The prediction of profits for all franchised stores is  
21 employed as example management data. It can be assumed that  
22 sales data for each store for each day is produced by a  
23 trend that affects overall sales for all franchised stores  
24 and sales factors that influence sales at each store.  
25 Example factors that influence sales at each store can be a  
26 change in the willingness of local consumers to purchase

1 goods that is due to advertisements disseminated by mass  
2 media, and advertisements prepared by and sales methods  
3 employed in each store. However, when analyzing management  
4 data, it is important that main factors, independent of the  
5 various factors referred to, be extracted that affect the  
6 overall cash flow for all franchised stores. Therefore,  
7 when we assume that the sales data for all franchised stores  
8 consist of an overall trend and the individual factors mixed  
9 by an unknown mixing matrix, the signal separation method of  
10 this embodiment, i.e., one type of independent component  
11 analysis method, can be effectively employed, so that the  
12 management analysis can be performed more quickly and  
13 accurately than by the conventional method.

14 The usage for the financial field can be considered in the  
15 same manner. For various portfolio products developed  
16 through financial engineering, portfolio return prediction  
17 can be performed for a set of several certificate issues,  
18 and a determination can be made as to how investment  
19 division should be handled. At this time, when the main  
20 element has been observed and designated that drives the  
21 setting of stock prices, an effective financial engineering  
22 product can be provided. Assuming that each stock price  
23 change can be modeled by the sum of several independent  
24 components, the independent component that drives the target  
25 setting can be obtained by employing the signal separation  
26 method of this embodiment. Then, when the trend of the  
27 independent component that largely affects the stock price  
28 is read and the estimated stock price is synthesized using  
29 the estimated mixing matrix, a more effective portfolio

1 return can be predicted. Further, when modeling is  
2 performed at a higher level, based on the above described  
3 idea, the accuracy of the independent component analysis  
4 method and the need for a fast convergence method are not  
5 lost. As is apparent from the fact that the independent  
6 component analysis can be correlated with the cost function,  
7 for which the high-order statistics is taken into account,  
8 it can be expected that the accuracy will be increased  
9 compared with using a combination of the principle  
10 components that are extracted using a method, such as a  
11 conventional principle component analysis method, for which  
12 the second-order statistics is taken into account. Further,  
13 in this embodiment, since only a short time is used to  
14 estimate a separation matrix, a separation matrix can be  
15 accurately estimated by using only a small amount of  
16 observed data, and a quick determination can be implemented.  
17 From this viewpoint, it is also apparent that the method of  
18 this embodiment can be effectively employed for the  
19 management/financial data analysis method in manners  
20 described and known to those skilled in the art.

21 As is described above, according to the invention, when  
22 multiple mixed signals are observed, only a small number of  
23 calculation steps are used to stably separate and extract  
24 the original signal from these multiple mixed signals.

25 The present invention can be realized in hardware, software,  
26 or a combination of hardware and software. It may be  
27 implemented as a method having steps to implement one or more  
28 functions of the invention, and/or it may be implemented as



1 an apparatus having components and/or means to implement one  
2 or more steps of a method of the invention described above  
3 and/or known to those skilled in the art. A visualization  
4 tool according to the present invention can be realized in a  
5 centralized fashion in one computer system, or in a  
6 distributed fashion where different elements are spread  
7 across several interconnected computer systems. Any kind of  
8 computer system - or other apparatus adapted for carrying out  
9 the methods and/or functions described herein - is suitable.  
10 A typical combination of hardware and software could be a  
11 general purpose computer system with a computer program that,  
12 when being loaded and executed, controls the computer system  
13 such that it carries out the methods described herein. The  
14 present invention can also be embedded in a computer program  
15 product, which comprises all the features enabling the  
16 implementation of the methods described herein, and which -  
17 when loaded in a computer system - is able to carry out these  
18 methods.

19 Computer program means or computer program in the present  
20 context include any expression, in any language, code or  
21 notation, of a set of instructions intended to cause a system  
22 having an information processing capability to perform a  
23 particular function either directly or after conversion to  
24 another language, code or notation, and/or after reproduction  
25 in a different material form.

26 Thus the invention includes an article of manufacture which  
27 comprises a computer usable medium having computer readable  
28 program code means embodied therein for causing one or more

functions described above. The computer readable program code means in the article of manufacture comprises computer readable program code means for causing a computer to effect the steps of a method of this invention. Similarly, the present invention may be implemented as a computer program product comprising a computer usable medium having computer readable program code means embodied therein for causing a function described above. The computer readable program code means in the computer program product comprising computer readable program code means for causing a computer to effect one or more functions of this invention. Furthermore, the present invention may be implemented as a program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for causing one or more functions of this invention.

It is noted that the foregoing has outlined some of the more pertinent objects and embodiments of the present invention. This invention may be used for many applications. Thus, although the description is made for particular arrangements and methods, the intent and concept of the invention is suitable and applicable to other arrangements and applications. It will be clear to those skilled in the art that modifications to the disclosed embodiments can be effected without departing from the spirit and scope of the invention. The described embodiments ought to be construed to be merely illustrative of some of the more prominent features and applications of the invention. Other beneficial results can be realized by applying the disclosed invention

1 in a different manner or modifying the invention in ways  
2 known to those familiar with the art.